

What is claimed is:

1. A computer database system for organizing, associating and retrieving information where characteristics regarding entities can be inferred from the characteristics of similar entities, having at least two sets of information located on one or more databases comprising a first and a second collection of information, wherein the first collection of information is joined by a chemical similarity join with the second collection of information to create a set of joined items, wherein the items are grouped together based on a plurality of properties.
2. The system of claim 1 wherein one or more properties of the plurality of properties is selected from the group consisting of chemical structure, synthesis pathway, binding data, biological activity, structure-activity relationship information, molecular weight, partition, coefficient, electric charge, size, efficacy, toxicology, manufacturer, price, and availability.
3. The system of claim 1 wherein the resultant joined items are reported to a user of the computer program.
4. The system of claim 3 wherein the user interacts with the joined items via remote communication link.
5. The system of claim 4 wherein the remote communication link is the Internet.

6. The system of claim 1 wherein the entities are biological compounds.
7. The system of claim 6 wherein the biological compounds are proteins.
8. The system of claim 1 wherein the entities are genes.